

STA 304 Paper 1

Sijian Wu

2/6/2022

Abstract

My abstract

#Data Cleaning Section

#Introduction

hhhhhh¹

#Data Source Column Description _id Unique row identifier for Open Data database

RSN This is the ID number for a building. The registration ID can be used to identify information pertaining to buildings in different RentSafeTO Open Data Sets.

YEAR_REGISTERED This is the year that the building first registered with RentSafeTO.

YEAR_EVALUATED

This represents the year of the building evaluation scores.

YEAR_BUILT

This is the year that the building was built in. Information is provided by the Building Owners/Managers.

PROPERTY_TYPE

This field informs users of whether a building is owed privately, by Toronto Community Housing Corporation (TCHC) or another assisted, social or supportive housing provider.

WARD

This is the ward that the building is located in. All data is provided based on the 25 ward system.

WARDNAME

This is the name of the ward

SITE_ADDRESS

This is the building address.

CONFIRMED_STOREYS

This is the number of storeys in a building.

CONFIRMED_UNITS This is the number of units in a building.

EVALUATION_COMPLETED_ON This is the date the evaluation was conducted on.

SCORE

This is the overall score of the building. The score is the sum total of each item that was evaluated. The formula to calculate scores is as follows: sum of all assigned scores during the evaluation / (number of unique items reviewed *5)

¹This footnote will appear at the bottom of the page.:<https://github.com/NnnnninaW/2022-STA304-1.git>

RESULTS_OF_SCORE

The score is used to determine whether an audit (which is a comprehensive examination of the building) takes place or whether another evaluation will be conducted in one, two or three years.

NO_OF_AREAS_EVALUATED

This is the number of items that were evaluated during a single evaluation.

ENTRANCE_LOBBY

This represents the condition of the entrance and/or lobby in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

ENTRANCE_DOORS_WINDOWS

This represents the condition of the entrance doors and windows in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

SECURITY

This represents the condition of the security system(s) in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

STAIRWELLS

This represents the condition of the stairwells in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

LAUNDRY_ROOMS

This represents the condition of the laundry room(s) in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

INTERNAL_GUARDS_HANDRAILS

This represents the condition of the internal guards and handrails in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

GARBAGE_CHUTE_ROOMS This represents the condition of the garbage/chute rooms in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

GARBAGE_BIN_STORAGE_AREA

This represents the condition of the garbage bin storage room or outdoor enclosure area. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

ELEVATORS

This represents the condition of the elevator(s) in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

STORAGE_AREAS_LOCKERS

This represents the condition of the storage areas/lockers in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

INTERIOR_WALL_CEILING_FLOOR This represents the condition of internal walls, ceilings and floors in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

INTERIOR_LIGHTING_LEVELS

This represents the condition of internal lighting levels in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

GRAFFITI

This score represents the severity of graffiti in a building. Scores range from 1 to 5, with 1 being a significant amount of graffiti and 5 being no graffiti.

EXTERIOR_CLADDING

This represents the condition of the exterior cladding/bricks/paint, flashing and drain pipes on a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

EXTERIOR_GROUNDS

This represents the condition of the exterior grounds of a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

EXTERIOR_WALKWAYS

This represents the condition of the exterior walkways of a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

BALCONY_GUARDS

This represents the condition of the balcony guards on a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

WATER_PEN_EXT_BLDG_ELEMENTS This represents the condition of water penetration of external elements of a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

PARKING_AREA

This represents the condition of the parking areas of a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

OTHER_FACILITIES

This represents the condition of other facilities in a building. Scores range from 1 to 5, with 1 being the lowest and 5 being the highest.

GRID

This is the grid that the building is located in. A grid represents a specific administrative area for bylaw enforcement.

LATITUDE

The latitude associated with the building address.

LONGITUDE

The longitude associated with the building address.

X

The projected X coordinate associated with the building address. The projected coordinate system is NAD27 MTM Zone 10.

Y

The projected Y coordinate associated with the building address. The projected coordinate system is NAD27 MTM Zone 10.